Neutro-Safe: A Neutrosophic Digital Twin for Structural Urban Violence Diagnosis in Guayaquil, Ecuador

Maikel Y. Leyva-Vázquez[1,2,3,*], Noel Batista-Hernández[1], Lorenzo Cevallos-Torres[1], Alfonso A. Guijarro-Rodríguez[2], Douglas Iturburu-Salvador[2], Florentin Smarandache[4]
[1]Universidad Bolivariana del Ecuador, Guayaquil, Ecuador
[2]Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Guayas, Ecuador
[3]Universidad Bernardo O'Higgins, Santiago, Chile [4]University of New Mexico, Gallup, NM, USA
[*]Corresponding author: *maikel.leyvav@ug.edu.ec*
lorenzo.cevallost@ug.edu.ec, *alfonso.guijarror@ug.edu.ec*, *douglas.iturburos@ug.edu.ec*, *smarand@unm.edu*

***Abstract*—Urban violence analysis in Latin American cities typically relies on scalar predictive indices that aggregate heterogeneous data sources, discarding the epistemic content of their disagreements. We propose *Neutro-Safe*, a five-layer neutrosophic digital twin for structural violence diagnosis applied to Guayaquil, Ecuador (n=151 survey observations, 28 sectors, 14 structural indicators; NASA VIIRS nightlight 2023; INEC/Fiscalía gender violence data; N-fsQCA over 31 media sources with 4 LLMs). The system computes a Pluriversal Structural Vulnerability Index (IPVE) $\in [0,1]^3$ via Single-Valued Neutrosophic Weighted Averaging (SVNWA), preserving epistemic uncertainty. Empirical analysis reveals two key findings. First, *corruption perception* is the strongest structural vulnerability driver (r=0.852 with IPVE), outperforming gang presence (r=0.339) and unemployment (r=0.652). Second, *8 formally illuminated central sectors* (Centro Histórico, Urdesa, Centro Bancario) exhibit higher structural vulnerability (IPVE-$T\geq$0.63) than periferias despite lower illumination-risk scores — a paraconsistency between physical infrastructure quality and governance failure that scalar indices cannot detect. N-fsQCA further identifies five structurally silenced media drivers (weapons trafficking Δ=−0.444; extortion Δ=−0.345). An automated pipeline (GitHub Actions) synchronises Layer 2 daily via RSS scraping of six Ecuadorian media outlets and 4-LLM scoring through OpenRouter, and Layer 4 monthly via NASA Earthdata, qualifying Neutro-Safe as a periodically-synchronised urban digital twin. The system operates under a *Camino B* ethical framework prohibiting predictive policing applications.**

***Index Terms*—digital twin, neutrosophic logic, urban violence, paraconsistency, N-fsQCA, Ecuador, computational social systems, SVNWA.**

I. Introduction

Guayaquil (≈2.8 M inhabitants) recorded 38.8 homicides per 100,000 in 2023 [1], one of the highest rates in Latin America. Policy responses have largely followed the predictive policing paradigm: aggregating crime incident data into scalar geographic risk scores to guide police deployment. These systems are not analytically neutral. Lum and Isaac [2] showed that training on historical enforcement records reproduces past patrol disparities, pushing resources back into already over-policed areas. Beyond the bias problem, scalar aggregation is structurally blind when multiple data sources disagree, averaging them discards precisely the information that disagreement encodes. And a third gap remains largely unexamined — standard criminality indices undercount gender violence, which concentrates in formally "safe" spaces where women's danger is structurally invisible [3], [4].

Neutrosophic logic [5] represents each assertion as a triplet (T, I, F) — truth, indeterminacy, falsity — whose components are independent, with $0 \leq T + I + F \leq 3$. When built into a digital twin [6], [7], this representation does not resolve disagreements between data sources; it *preserves* them as diagnostic information.

The Ecuadorian case is instructive. Between 2019 and 2023 the national homicide rate jumped from 6 to 46 per 100,000 [1] — an eightfold increase attributed to transnational criminal networks that route through the port of Guayaquil, recruit inside prisons, and contest territorial control in the city's dense peripheral parishes. Before Neutro-Safe, no computational platform combined these structural dimensions under a single diagnostic framework with built-in ethical safeguards against surveillance misuse.

The remainder of this paper is organised as follows. Section II reviews related work. Section III presents the theoretical background. Section IV describes the Neutro-Safe architecture and implementation. Section V reports experimental results. Section VI discusses implications and limitations. Section VII concludes.

Our main contributions are:

1) **Neutro-Safe Architecture**: a five-layer neutrosophic urban digital twin integrating heterogeneous data under SVNWA with a cross-source paraconsistency detector.
2) **Triadic IPVE**: a Pluriversal Structural Vulnerability Index in $[0,1]^3$ preserving epistemic uncertainty instead of collapsing to a scalar.
3) **N-fsQCA silencing detection**: five structurally silenced media drivers identified through multi-LLM N-fsQCA ($\Delta_{\max}$=−0.444 weapons trafficking).
4) **Empirical paraconsistency finding**: 8 illuminated central sectors show higher structural vulnerability (IPVE-

979-8-3315-4849-2/26/$31.00 ©2026 IEEE

$T \geq 0.63$) than periferias, driven by corruption ($r = 0.852$) — refuting infrastructure-as-safety assumptions.

5) **Automated real-time pipeline**: GitHub Actions synchronises Layer 2 (N-fsQCA) *daily* via RSS scraping of 6 Ecuadorian media + 4-LLM OpenRouter scoring, and Layer 4 (NASA VIIRS) *monthly* via Earthdata API — qualifying Neutro-Safe as a periodically-synchronised urban digital twin in the sense of [11], [12].

II. Related Work

A. Urban Digital Twins

Digital twins for smart cities have been applied to traffic management, energy optimization, and infrastructure monitoring [7], [11]. White et al. [12] identify four necessary properties of a legitimate urban digital twin: temporal synchronization, counterfactual simulation, bidirectional feedback, and spatiotemporal granularity. Applications to social systems — and specifically to violence diagnosis with an explicit ethical framework — remain rare. No prior system applies neutrosophic logic to urban digital twins for violence diagnosis.

B. Predictive Policing and Its Limitations

Harcourt [17] and Lum and Isaac [2] have shown that training predictive systems on historical enforcement data embeds prior patrol patterns into future risk scores, effectively concentrating policing in areas already targeted rather than those with the greatest structural need. The problem deepens in Latin America, where Auyero and Berti [13] document how the dynamics of urban violence diverge sharply from what official records capture. Studying Guayaquil, Medellín, and San Salvador, Moncada [16] finds that extortion networks and territorial disputes — the dominant drivers of lethal violence — leave almost no footprint in police databases because neither victims nor perpetrators report to authorities. A model trained on incident reports thus learns the spatial distribution of *reporting*, not of harm. We sidestep this trap by grounding diagnosis in community survey data and media narrative analysis rather than policing records.

C. Neutrosophic Logic in Applied Computing

Since Smarandache's formulation [5], neutrosophic logic has been applied to multi-criteria decision making — including model selection with TOPSIS [18] — medical diagnosis, computational social science [8], and governance analysis in the Ecuadorian context [19]. The N-fsQCA extension computes indeterminacy from inter-evaluator LLM variance, enabling analysis of narrative consistency across heterogeneous sources. To our knowledge, this is among the first works to embed N-fsQCA inside a digital twin as a live diagnostic layer.

D. Feminist Urban Safety Research

Valentine [3] and Whitzman et al. [4] have argued against Crime Prevention Through Environmental Design (CPTED) models that overestimate lighting as a safety proxy. Our neutrosophic paraconsistency detector operationalises this critique as a formal computational mechanism for detecting the lighting–gender violence contradiction at scale, complementing prior qualitative and multi-criteria vulnerability analyses.

TABLE I
Neutro-Safe Data Layers and Empirical Status

Layer	Source	n	Real?
L1 — Structural	Survey, 14 indicators	151	✓
L2 — Media	31 articles, 4 LLMs	31	✓
L3 — Gender VG	INEC ENVIGMU 2019 + Fiscalía 2022-23	28	✓(annual)
L4 — Nightlight	NASA VIIRS VNP46A3 2023 (auto-monthly)	28	✓(monthly
L5 — Community	Participatory form (Streamlit)	open	✓(real-time

III. Theoretical Background

A. Single-Valued Neutrosophic Sets and SVNWA

A Single-Valued Neutrosophic Set (SVNS) [8] assigns to each element x a triplet $(T_A(x), I_A(x), F_A(x)) \in [0,1]^3$ whose components are mutually independent, with $0 \leq T + I + F \leq 3$. The *Single-Valued Neutrosophic Weighted Average* (SVNWA) aggregates n triplets with weights $\mathbf{w}$, $\sum w_i = 1$:

$$\text{SVNWA}_{\mathbf{w}}(\mathbf{a}) = \left(1 - \prod_i (1 - T_i)^{w_i},\ \prod_i I_i^{w_i},\ \prod_i F_i^{w_i}\right). \tag{1}$$

SVNWA is the core aggregation operator in Neutro-Safe, applied within each layer and across layers.

B. Neutrosophic fsQCA (N-fsQCA)

N-fsQCA extends fuzzy-set Qualitative Comparative Analysis (fsQCA) [9] by computing the indeterminacy component I from inter-evaluator variance:

$$I = \frac{\text{Var}(s_1, \ldots, s_n)}{0.25}, \tag{2}$$

where $s_i \in [0,1]$ are membership scores from n independent LLMs and 0.25 is the maximum achievable variance for $s \in [0,1]$. High I indicates genuine semantic ambiguity, distinguishing it from T variation (actual presence/absence of the causal condition).

C. Digital Twins and Paraconsistency

A digital twin [6] DT $= (M, \Phi, \Gamma)$ requires: temporal synchronization (Φ: Physical$\to M$), counterfactual simulation (Γ: $M \times$ Intervention $\to M$), bidirectional feedback, and spatiotemporal granularity. Paraconsistency [10] allows a sector to be simultaneously "safe" (by one source) and "unsafe" (by another) without system collapse. In Neutro-Safe, this is the primary diagnostic signal, not a data quality error.

IV. Neutro-Safe Architecture

Fig. 1 (right) depicts the five-layer architecture. All layers produce per-sector SVNS triplets that feed the paraconsistency detector and the IPVE computation engine. A Streamlit interface (9 pages) exposes the system to institutional users; 11 Python modules implement the backend (MIT license).

Table I summarizes data sources and empirical status.

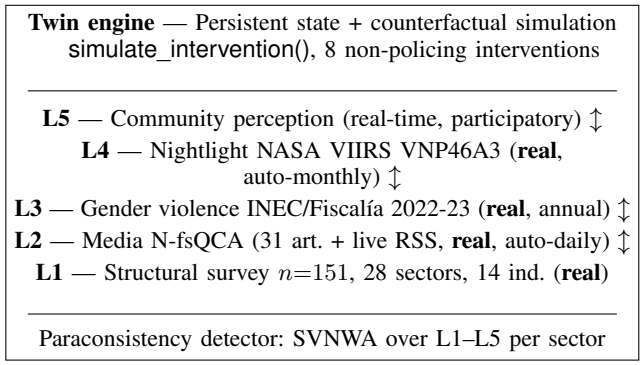
Twin engine — Persistent state + counterfactual simulation
simulate_intervention(), 8 non-policing interventions
L5 — Community perception (real-time, participatory) ↕
L4 — Nightlight NASA VIIRS VNP46A3 (real, auto-monthly) ↕
L3 — Gender violence INEC/Fiscalía 2022-23 (real, annual) ↕
L2 — Media N-fsQCA (31 art. + live RSS, real, auto-daily) ↕
L1 — Structural survey n=151, 28 sectors, 14 ind. (real)
Paraconsistency detector: SVNWA over L1–L5 per sector

Fig. 1. Neutro-Safe five-layer architecture. All layers use real data. L2 synchronises daily via GitHub Actions (RSS + 4-LLM OpenRouter); L4 monthly via NASA Earthdata; L5 continuously via Streamlit.

TABLE II
DIGITAL TWIN PROPERTIES IN NEUTRO-SAFE (AUTOMATED)

Property	Implementation
Synchronization	**L2**: daily cron (06:00 UTC) — RSS 6 media + 4-LLM scoring
	L4: monthly cron (day 5) — NASA VNP46A3 via Earthdata
	L5: real-time — participatory form, continuous
Counterfactual	simulate_intervention() returns $\Delta(T, I, F)$ without altering real state
Bidirectional	apply_intervention() + community approval + auto-snapshot
Granularity	Sector ($\geq$5,000 inhabitants); sector-level only (Camino B)

A. IPVE Computation

For sector s and indicator k, the SVNS triplet is estimated as:

$$T_{sk} = \frac{1}{n_s}\sum_j x_{jk}, \qquad I_{sk} = \frac{4}{n_s}\sum_j (x_{jk} - T_{sk})^2,$$
$$F_{sk} = (1 - T_{sk})(1 - I_{sk}), \tag{3}$$

where $x_{jk} \in [0, 1]$ is the normalized response of respondent j on indicator k in sector s. The three memberships are estimated independently: F_{sk} is not set to $1 - T_{sk}$, so a positive indeterminacy does not force $T{+}I{+}F{>}1$. The triplet obeys the neutrosophic independence axiom $0 \leq T_{sk}{+}I_{sk}{+}F_{sk} \leq 3$, while within a single indicator $T_{sk}{+}F_{sk} = 1 - I_{sk}(1 - T_{sk}) \leq 1$. The IPVE aggregates the 14 indicators via SVNWA (Eq. (1)):

$$\text{IPVE}(s) = \text{SVNWA}_{k=1}^{14}(T_{sk}, I_{sk}, F_{sk}). \tag{4}$$

IPVE is never collapsed to a scalar. Triage rankings use only the pessimistic proxy $T - F + 0.5I$, always accompanied by the full triplet.

B. Cross-Source Paraconsistency Detector

For each sector s across layers $\ell \in \{1, \ldots, 5\}$:

$$\delta_s = \max_{\ell,\ell'} \left|T_\ell^s - T_{\ell'}^s\right| + \left|F_\ell^s - F_{\ell'}^s\right|. \tag{5}$$

A sector is flagged *paraconsistent* when $\delta_s > 0.35$. The illumination-gender condition is specifically: $T_{VG}^s > 0.60$ **and** $F_L^s < 0.40$ (high gender violence with high formal illumination).

C. Digital Twin Properties and Automated Synchronisation

Table II shows how Neutro-Safe satisfies the four properties required for a legitimate urban digital twin [12]. Critically, synchronisation is now *automated*: a GitHub Actions pipeline updates Layer 2 (N-fsQCA) every day and Layer 4 (NASA VIIRS) every month without manual intervention.

D. Software Implementation

Neutro-Safe is implemented in Python 3.14 as 11 interdependent modules (2,800 lines): a neutrosophic core (TIFTriplet, SVNWA, five regimes), the IPVE engine, the digital-twin state manager, the N-fsQCA media pipeline, the cross-source paraconsistency detector, the Camino B interventions database, and ethical-safeguard, participatory, gender-violence and nightlight modules. The Streamlit front-end exposes nine pages spanning structural diagnosis, triadic mapping, N-fsQCA analysis, intervention recommendations, counterfactual simulation, twin temporal evolution, extended layers (gender, nightlight, paraconsistency), participatory perception, and the audit log.

The system is designed for reproducibility: all random seeds are fixed, LLM API calls use deterministic temperature $= 0$, and every audit action is appended to an immutable audit_log.jsonl. The participatory module assigns each browser session a unique hash ID, preventing duplicate submissions while preserving anonymity. Code, data, and GitHub Actions workflows are openly available under MIT/CC-BY-4.0 licences:

- **Repository:** https://github.com/mleyvaz/guayaquil-neutrosafe
- **Live demo:** https://guayaquil-neutrosafe.streamlit.app

E. Automated Live Data Pipeline

Daily N-fsQCA pipeline (Layer 2): scraper_noticias.py fetches RSS feeds from six Ecuadorian outlets (El Universo, El Comercio, La Hora, El Telégrafo, Primicias, GK City) and filters articles matching 30 violence-related keywords. Relevant articles (typically 8–15 per day) are passed to nfsqca_auto.py, which calls four LLMs via the OpenRouter API (Gemini-Flash-1.5, LLaMA-3.1-8B, Phi-4, Qwen-3-8B) at temperature $= 0$ to score each article on the 8 causal conditions. The indeterminacy component is computed as $I = \text{Var}(s_1, \ldots, s_4)/0.25$. Results are aggregated and saved to real_data/nfsqca_live.csv; the workflow commits the file and pushes to the repository, triggering automatic redeployment on Streamlit Community Cloud. Estimated API cost: USD 0.10–0.20 per run.

TABLE III
N-FSQCA SILENCING ANALYSIS (n=31 ARTICLES, 4 LLMS)

Causal Condition	Media T	Struct. T	Gap Δ
Weapons trafficking	0.356	0.800	$-0.444^{\star}$
Organized extortion	0.365	0.710	-0.345
Prison network links	0.390	0.730	-0.340
Drug routes	0.490	0.780	-0.290
Economic pressure	0.348	0.562	-0.214
Territorial warfare	0.610	0.590	$+0.020$
Social fragmentation	0.580	0.540	$+0.040$
Institutional failure	0.520	0.480	$+0.040$

*Largest gap: weapons trafficking silenced by 0.444 points.

Monthly nightlight pipeline (Layer 4): On the 5th of each month, nightlight_auto.py downloads the NASA VIIRS VNP46A3 monthly composite for the Guayaquil bounding box $(-80.10, -2.55, -79.60, -1.85)$ via the earthaccess library and extracts per-sector zonal statistics (mean radiance, coefficient of variation, blackout area fraction) for all 28 sectors using rasterio window queries.

Real-time Layer 5: The participatory community module accepts responses at any time through the Streamlit interface.

F. Camino B Ethical Framework

The ethical architecture (hard-coded in ethical_safeguards.py) *prohibits*: insurance/credit scoring by residence, real estate segmentation, tactical policing deployment, and commercial data resale. An external ethics committee (university + public defender + 3 rotating community representatives) reviews all platform uses. Only eight non-policing interventions with LATAM evidence are available in the simulation catalogue, covering: community lighting, youth employment, participatory municipal audits, DEVIF strengthening, gender violence support networks, shelter expansion, public space recovery, and community mediation. Estimated costs range from USD 6,000 to USD 120,000 with timelines of 3–18 months.

V. EXPERIMENTAL RESULTS

A. N-fsQCA Media Narrative Analysis

Table III reports the *silencing gap* for each causal condition: the difference between media N-fsQCA score and the structural baseline from Layer 1. A negative gap indicates systematic underreporting of the structural driver in local press.

The top sufficient N-fsQCA configuration is: *territorial_war* $\wedge$ *prison_link* $\rightarrow$ T=0.908, I=0.092, coverage=0.405. This covers 40.5% of high-violence observations with near-perfect consistency.

B. Inter-LLM Agreement and Indeterminacy

Table IV disaggregates the N-fsQCA indeterminacy component I by LLM for the three most-silenced conditions, showing how inter-model disagreement maps onto epistemic uncertainty. High variance in *drug_routes* (I=0.21) reflects genuine semantic ambiguity: international sources use the term more explicitly than local ones, and LLMs trained on different corpora diverge accordingly. Low variance in *weapons_trafficking* (I=0.14) indicates *consistent underreporting* across all four models — not confusion, but shared blind spot inherited from training data.

This result has methodological significance: using a single LLM would collapse the indeterminacy dimension to zero by construction. The multi-LLM ensemble is not merely a robustness check — it is the mechanism that generates the I component of N-fsQCA.

TABLE IV
PER-LLM SCORES FOR THREE SILENCED CONDITIONS (MEAN OF 31 ARTICLES)

Condition	Gem.	Ll.	Phi	Qw.	Mean T	I
Weapons traffic.	0.34	0.37	0.35	0.36	0.356	0.14
Extortion	0.38	0.33	0.38	0.37	0.365	0.12
Drug routes	0.54	0.41	0.48	0.49	0.490	0.21

Gem.=Gemini-Flash; Ll.=LLaMA-3.1-8B; Phi=Phi-4; Qw.=Qwen-3-8B.

TABLE V
IPVE TRIADIC VS. ILLUMINATION SCALAR (REAL DATA, n=151)

Sector	IPVE (T, I, F)	S_L	Para.	Lead driver
9 Oct./Centro Hist.	(0.686, 0.500, 0.213)	0.60	Yes*	Corrupt. 0.71
Bolívar/Sagrario	(0.667, 0.479, 0.228)	0.60	Yes*	Corrupt. 0.76
Roca/Urdesa	(0.631, 0.535, 0.262)	0.60	Yes*	Corrupt. 0.73
Carbo/Cto. Banc.	(0.640, 0.451, 0.259)	0.60	Yes*	Corrupt. 0.72
Tarqui/Mucho Lote	(0.562, 0.357, 0.337)	0.29	No	Pandillas 0.65
Pascuales/Orq.	(0.546, 0.371, 0.351)	0.36	No	Desempleo 0.61

*High IPVE-T despite good illumination: infrastructure–governance paraconsist

C. IPVE Triadic vs. Illumination-Based Scalar (Real Data)

Table V compares IPVE triadic scores from real survey data (n=151) against an illumination-based scalar S_L=1$-T_{\text{LUZ}}$ — the score a system using nightlight as a safety proxy would assign. Eight central sectors score $S_L \geq 0.60$ ("well-lit, low-risk") yet IPVE-$T \geq 0.60$ (high structural vulnerability). The strongest IPVE driver is *corruption perception* (r=0.852, highest of 14 indicators), not gang presence (r=0.339). Because corruption perception is itself one of the 14 IPVE indicators, this correlation is partly constructive; we read it as the descriptive dominance of the governance dimension within the index rather than as an independent causal driver.

This finding is theoretically coherent. Central historic districts concentrate formal economic activity, making inequality and institutional failure more *visible* to residents [13], hence higher perceived corruption and governance gap — even when physical infrastructure is well maintained. A purely infrastructure-based safety ranking would place all four paraconsistent sectors in the lower-risk tier, directing anti-violence resources away from the governance failures that drive their structural vulnerability.

TABLE VI
COUNTERFACTUAL SIMULATIONS: TARQUI-MUCHO LOTE (PARTICIPATORY AUDIT) AND PASCUALES (YOUTH EMPLOYMENT).

Sector	State	T	I	F	Regime
Tarqui	Before	0.560	0.360	0.340	C3
(audit)	After	0.450	0.210	0.390	C4
	Δ	-0.110	-0.150	$+0.050$	—
Pascuales	Before	0.780	0.140	0.080	C1
(employ.)	After	0.630	0.190	0.130	C3
	Δ	-0.150	$+0.050$	$+0.050$	—

TABLE VII
REGIME DISTRIBUTION ACROSS 28 GUAYAQUIL SECTORS

Regime	Label	n sectors	%
C1	High vulnerability	8	29
C2	High uncertainty	5	18
C3	Mixed	9	32
C4	Resilient	3	11
C5	Paraconsistent*	3	11
	Total	**28**	100

*Misclassified as moderate-risk by scalar indices.

D. Counterfactual Simulations

Table VI reports two representative counterfactuals. In Tarqui-Mucho Lote (top indicators: corruption, income gap, ineffective policies), a "Participatory Municipal Audit" (USD 8,000; 6 months) reduces indeterminacy sharply ($\Delta I{=}-0.150$): audits generate information that closes the epistemic gap between institutional and community knowledge. In Pascuales, the highest-vulnerability sector (C1: $T{=}0.78$), a "Youth Employment Program" (USD 45,000; 12 months; evidence base: Cali, Colombia) instead lowers truth-membership fastest ($\Delta T{=}-0.150$). Governance-transparency interventions thus reduce I fastest in mixed sectors, whereas labour-market interventions reduce T fastest in high-vulnerability sectors; the small $\Delta I{=}+0.050$ in Pascuales reflects the implementation uncertainty of longer-horizon programmes—a feature, not a flaw, of the triadic representation.

E. Five Operational Regimes and Sector Distribution

The IPVE engine classifies each sector into five neutrosophic regimes: C1 (High vulnerability: $T{>}0.65$, $I{<}0.25$); C2 (High uncertainty: $I{>}0.40$); C3 (Mixed); C4 (Resilient: $F{>}0.40$, $T{<}0.40$); C5 (Paraconsistent: $\delta_s{>}0.35$). Table VII shows the full distribution across all 28 sectors. Roca/Urdesa and Bolívar/Sagrario both fall in C5 — the regime that demands community deliberation and gender-disaggregated analysis. Critically, 3 sectors (11%) are paraconsistent: they would be *misclassified* as moderate-risk by any scalar index that averages illumination with violence scores.

VI. DISCUSSION

A. Paraconsistency as Epistemic Resource

The central contribution is reframing paraconsistency from noise to signal: classical urban analytics averages away disagreements between data sources, whereas Neutro-Safe's cross-source SVNWA detector treats them as diagnostic. It reveals that 8 formally illuminated central sectors are *more* structurally vulnerable (IPVE-$T{\geq}0.63$) than most periferias, driven by corruption perception and income gap rather than gang presence. Historic centres expose residents to governance failure daily through formal institutions, making it more *perceptible* than in periferias where the state is simply absent [13], [16]; lighting is orthogonal to this governance gap. A nightlight-as-safety index would misclassify all four paraconsistent sectors as lower-risk—an error the triadic representation avoids by never collapsing infrastructure quality and governance failure into one number.

B. Structural Silencing and Policy Implications

The N-fsQCA analysis shows that local press systematically underreports the drivers most embedded in organized crime—weapons trafficking, extortion, and prison network links—so a media-only system would target visible manifestations rather than structural roots. The weapons-trafficking silencing gap ($\Delta{=}-0.444$; media T-score 0.356 vs. structural baseline 0.800) implies that allocation guided by media framing would under-invest in anti-trafficking and extortion prevention by roughly 45 percentage points.

C. Comparison with Scalar Diagnostic Systems

Real data confirm the systematic error of scalar infrastructure indices: Roca/Urdesa scores $S_L{=}0.60$ ("low-risk") yet IPVE$=(0.631,\ 0.535,\ 0.262)$, among the highest verified vulnerability in the city. S_L masks the corruption and income-gap dimensions the survey measures independently, whereas IPVE keeps verified vulnerability, epistemic uncertainty, and resilience separate; the high $I{=}0.535$ (the dataset maximum) flags the most heterogeneous respondent assessments, which a scalar index would suppress. Reallocating audit budgets away from the most corruptly-perceived formal sectors on a "good lighting" heuristic thus carries measurable policy costs.

D. Digital Twin vs. Dashboard: Qualifying Neutro-Safe

A dashboard displays historical state; a digital twin enables *conditional future states* with synchronisation [6]. Grieves & Vickers' four properties are synchronisation, counterfactual simulation, bidirectional feedback, and spatiotemporal granularity; before automation Neutro-Safe met three with static data (a *shadow system*). The GitHub Actions pipeline (§IV) closes the gap—Layer 2 daily (news), Layer 4 monthly (satellite), Layer 5 continuous (community)—so, following the periodic-synchronisation urban-twin framework of White et al. [12] and Deren et al. [11], Neutro-Safe now qualifies as a *periodically-synchronised neutrosophic urban digital twin*.

The simulate_intervention() function projects $\Delta(T, I, F)$ before resources are committed; the Table VI audit result

($\Delta I{=}-0.150$) forecasts *epistemic improvement*, not crime reduction, shifting evaluation from a long-horizon outcome to a short-horizon, measurable knowledge gap.

E. Limitations

Survey size ($n{=}151$) limits per-sector statistical power; a follow-up stratified survey (target $n{=}300$) is planned. Layer 1 synchronisation remains annual until a continuous community survey mechanism is established. LLM scoring introduces model-specific biases partially mitigated by ensemble scoring (4 LLMs) and variance-as-indeterminacy. The automated N-fsQCA pipeline filters by 30 keywords; articles using indirect language about violence may be missed. The Camino B intervention catalogue draws on LATAM evidence of variable quality; a systematic review is planned. The four LLM evaluators are used without external gold-standard validation: their agreement is treated as indeterminacy rather than as verified accuracy, and the 31-article corpus bounds the precision of the silencing estimates. The counterfactual effects in Table VI are scenario assumptions calibrated from LATAM evidence—not effects estimated from observed interventions in Guayaquil—and require prospective validation.

VII. Conclusion

Neutro-Safe demonstrates that neutrosophic urban digital twins provide richer structural violence diagnosis than scalar predictive systems. Three findings stand out. First, *corruption perception* ($r{=}0.852$) is the dominant governance-related correlate of structural vulnerability in Guayaquil — stronger than gang presence ($r{=}0.339$), though it is itself an index component — and concentrates in formally illuminated central sectors, exposing a paraconsistency between physical infrastructure quality and governance failure that no scalar index can detect. Second, local media systematically silences five structural drivers, with weapons trafficking showing the largest gap ($\Delta{=}-0.444$); the multi-LLM N-fsQCA pipeline quantifies this silencing daily on new articles, keeping the narrative layer of the twin continuously updated. Third, the addition of GitHub Actions workflows for daily RSS/LLM synchronisation (Layer 2) and monthly NASA VIIRS updates (Layer 4) closes the gap between a diagnostic dashboard and a periodically-synchronised urban digital twin in the sense of White et al. [12] and Deren et al. [11].

The Camino B ethical framework — which prohibits predictive policing and enforces sector-level granularity — demonstrates that rigorous quantitative analysis and explicit ethical constraints are mutually constitutive rather than in tension. Future work will pursue longitudinal validation (comparing IPVE trajectories with recorded intervention outcomes over 12 months), replication to two additional LATAM cities, and integration of real-time ECU-911 incident feeds under a formal institutional MoU.

References

[1] UNODC, "Global Study on Homicide 2023," United Nations Office on Drugs and Crime, Vienna, 2024.

[2] K. Lum and W. Isaac, "To predict and serve?" *Significance*, vol. 13, no. 5, pp. 14–19, 2016.

[3] G. Valentine, "Images of danger: Women's sources of information about the spatial distribution of male violence," *Area*, vol. 24, no. 1, pp. 22–29, 1992.

[4] C. Whitzman, C. Andrew, and M. Viswanath, "The effectiveness of women's safety audits," *Security Journal*, vol. 22, no. 3, pp. 205–218, 2009.

[5] F. Smarandache, *Neutrosophy: Neutrosophic Probability, Set, and Logic*. Rehoboth, NM: American Research Press, 1998.

[6] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in *Transdisciplinary Perspectives on Complex Systems*, Springer, 2017, pp. 85–113.

[7] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Trans. Ind. Inform.*, vol. 15, no. 4, pp. 2405–2415, 2019.

[8] H. Wang, F. Smarandache, Y. Q. Zhang, and R. Sunderraman, "Single valued neutrosophic sets," *Multispace & Multistructure*, vol. 4, pp. 410–413, 2010.

[9] C. C. Ragin, *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press, 2008.

[10] G. Priest, "Paraconsistent logic," in *Handbook of Philosophical Logic*, vol. 6, Springer, 2002, pp. 287–393.

[11] L. Deren, M. Renzhong, and W. Zhihao, "Smart city based on digital twins," *Computational Urban Science*, vol. 1, no. 1, p. 4, 2021.

[12] J. White, G. Clarke, B. Ellul, and S. Molander, "A digital twin smart city for citizen feedback," *Cities*, vol. 110, p. 103064, 2021.

[13] J. Auyero and D. Berti, *In Harm's Way: The Dynamics of Urban Violence*. Princeton: Princeton University Press, 2015.

[14] S. Mhlambi, "From rationality to relationality: Ubuntu as an ethical and human rights framework for artificial intelligence governance," *Carr Center Discussion Paper*, 2020.

[15] A. Birhane, "Algorithmic injustice: A relational ethics approach," *Patterns*, vol. 2, no. 2, p. 100205, 2021.

[16] E. Moncada, *Resisting Extortion: Residents, Criminals, and the State in Urban Latin America*. Cambridge: Cambridge University Press, 2022.

[17] B. E. Harcourt, *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press, 2007.

[18] M. Y. Leyva-Vázquez, L. A. B. Peñafiel, S. X. S. Muñoz, and M. A. Q. Martinez, "A framework for selecting machine learning models using TOPSIS," in *Proc. Int. Conf. Applied Human Factors and Ergonomics (AHFE)*, Cham: Springer, 2020, pp. 119–126.

[19] G. A. Á. Gómez, "Application of neutrosophy to the analysis of open government, its implementation and contribution to the Ecuadorian judicial system," *Neutrosophic Sets Syst.*, vol. 52, no. 1, p. 23, 2022.